REMARKS/ARGUMENTS

Reconsideration of this application is respectfully requested.

Specification

The abstract of the disclosure was objected to because it was longer than 150 words. The abstract is amended to comply with MPEP § 608.01(b). The objection is thereby traversed.

Claim Rejection -35 U.S.C. § 112

Claim 20 was rejected under 35 U.S.C. § 112, second paragraph, because due to a typographical error claim 20 depended from itself. The dependency of claim 20 has been corrected and it now depends from claim 19, as intended. The rejection of claim 20 is thereby traversed.

Claim Rejection -35 U.S.C. § 102

The Office Action rejected claims 1-19 and 21-28 under 35 U.S.C. § 102(b) as being anticipated by Burrows in United States Patent No. 6,745,898.

Burrows teaches a method for generating a compressed index of information of records of a database in which a compressed index of information is generated. The information is stored as plurality of records in a database. Indexable portions of information are sequentially parsed to generate words and metawords. The words represent the portions of information, and the metawords represents attributes of the portions of information. A location is sequentially assigned to each word and metaword in the order that the portions are parsed to form pairs. The pairs are sorted first according to the words and metawords, and second according to the locations. Index entries are written to a memory for each unique word and metaword. Each index entry includes a word entry or a metaword entry, and one or more location entries. The word and metaword entries use a prefix encoding which indicates the number of bytes that the unique word or metaword of a next index entry has in common with the unique word or metaword of a previous index entry. The location entries used a delta value encoding..

In accordance with Burrows, each **word** on a page is assigned a sequential number. The locations of the words indicate the relative order in which the parsing module identified the words in the pages. Each page has a first word and last word. The first word of the very first page parsed has an associated location "1", the next word has a location "2" and so on. From the perspective of the parsing module, the first word of a next page is considered to be positionally adjacent to the last word of a previous page (column 6, lines 25-44).

As noted above, Burrows uses a delta value encoding. The delta value means that the location is expressed as a relative offset in locations from a previous location. The first location for a particular word can be the offset from location "0". For example if a first occurrence of the word "the" is at location "100", and next occurrences are at locations "130" and "135", the delta values are respectively expressed as 100, 30 and 5 (column 10, lines 25-31).

Burrows therefore teaches that the locations of words and metawords in all pages parsed are stored using delta value encoding. The claims of the instant invention have been amended to clearly distinguish over the teachings of Burrows.

Claims 1, 13, 18, 22 and 25 have been canceled from the application.

Claim 2 is amended to claim a method of compressing file-reference data related to information symbols in a source file, comprising steps of: retrieving location data related to each location of respective ones of the information symbols in the source file; encoding the location data in the encoded data structure as one or more runs, each run including a first code for representing a line number of a first location of the information symbol in the source file encoded by the run, and, if the information symbol occurs more than once in the section of the source file, a second code providing a bit map representing one or more additional lines offset from the line number in the source file in which the information symbol occurs at least once in the section of the source file; compressing the

location data using a run encoding compression method to construct an encoded data structure; and storing the encoded data structure in a computer-readable storage medium.

Claim 2 as amended clearly distinguishes over Burrows because the location of information symbols in the source file is stored as a first code for representing a line number of a line including at least one first occurrence of the symbol and a second code providing a bit map representing one or more additional lines offset from the line number in the source file in which the information signal occurs at least once in the section of the source file.

Burrows neither teaches nor suggests either of these methods of encoding the locational data in the encoded data structure. As noted above, Burrows teaches encoding the position of all words and metawords. Burrows encodes the position of every occurrence of each word, regardless of how many times it appears on a page or a line. Burrows does not teach anything about differentiating lines or indexing a line only once, regardless of how many instances of the information symbol appear in the line. Furthermore, Burrows fails to teach or suggest providing a second code providing a bit map representing additional lines offset from the line number in the source code of the first occurrence. In contrast, Burrows uses delta encoding, in which whole numbers are used to encode relative offsets.

It is therefore respectfully submitted that claim 2 clearly distinguishes over the teachings of Burrows and the rejection of claims 1 and 2 is traversed.

Claim 3 is amended to remove redundant subject matter in view of the amendment to claim 2 and now claims that the first code is a binary number. Claim 4 is likewise amended to remove redundant subject matter in view of the amendments to claim 2 and now claims that the second code is a binary string that serves as the bit map. Claims 7 and 10-12 are amended to depend from amended claims 2.

In view of the amendment of claim 2, and for reasons set forth above, the rejection of claims 3-12 is likewise traversed.

As noted above, claim 13 is cancelled and the subject matter thereof is incorporated in claim 14. Claim 14 now claims a computer-readable medium containing a file-reference data structure, comprising one or more distinct information symbols and compressed file-reference data representing one or more locations of respective ones of the information symbols in a source file, the compressed file-reference data comprising run encoded location data generated by a run encoding compression method, an encoded data structure comprising one or more runs, each run comprising a first code for representing a line number of a first location of an information symbol in a section of the source file encoded by the run, and, if the information symbol occurs more than once in the source file, a second code comprising a bitmap representing one or more additional locations of the information symbol in the section of the source file.

Amended claim 14 includes all the limitations of amended claim 2 and for reasons set forth above in detail, claim 14 now clearly distinguishes over Burrows. The rejection of claims 14-17 is thereby traversed.

As noted above, claim 18 is cancelled, and the subject matter thereof incorporated in claim 14. Claim 19 is amended to improve clarity. Claim 21 is amended to depend from amended claim 14. The rejection of claims 18, 19, and 21 is therefore likewise traversed.

As noted above, claim 22 is cancelled and the subject matter thereof incorporated in amended claim 24.

Claim 24 now claims an apparatus for compressing file-reference data related to information symbols in a source file, comprising: means for retrieving location data related to each location of respective ones of the information symbols in the source file; means for compressing the location data using a run encoding compression algorithm to construct an encoded data structure, wherein the run encoding compression algorithm is adapted to analyze the location data and construct at least one run associated with each information symbol, each run comprising at least a first code indicating a line number in

the source file in which the information symbol appears and a binary string that serves as a bitmap offset from the first code to indicate a line location of additional occurrences of the information symbol in the source file; and means for storing the encoded data structure in a computer-readable storage medium.

Claim 24 is thereby amended to include all the limitations of amended claim 2 and for reasons set forth above likewise clearly distinguishes over the teaches of Burrows. The rejection of claim 24 is thereby traversed.

Claim 23 is amended to depend from amended claim 24. Claim 25 is cancelled. Claim 26 is amended to depend from amended claim 24 in view of the cancellation of claim 25. The rejection of claims 22-28 is thereby traversed.

In view of the above-noted amendments to the claims, and for reasons set for the above in detail, this application is now considered to be in a condition for allowance. Favourable reconsideration and early issuance of a Notice of Allowance are therefore requested.

Respectfully submitted,

By _____
Max R. Wood
Registration No.40,388
Attorney for Applicant

MRW/ma
Address: OGILVY RENAULT
1600-1981 McGill College Avenue
Montreal, QC H3A 2Y3
Canada